Efficient Generation of K-point Grids for Quantum ESPRESSO and Catalytic Role of CoP (101) for Electrochemical Hydrogen Evolution Compared with Pt (100) and Pt (111)

By

Wan Wan

A thesis submitted to Johns Hopkins University in conformity with the requirements for the degree of Master of Engineering

Baltimore, Maryland

May, 2018

Abstract

In computational material area, various properties of materials are determined by integrals over the Brillouin zone in reciprocal space. K-points are a set of discrete points to approximate these integrals. Our group developed and implemented called a k-point server to provide the fewest irreducible k-point grids for a required level of accuracy. Before this paper, our group had had implemented our grids into VASP and found out the computational cost of using this method is twice as low as the conventional methods for generating Monkhorst-Pack grids. In this paper, we applied this implement in Quantum ESPRESSO and tested its efficiency. Benchmark results on several randomly selected materials indicate that for well-converged calculations in Quantum ESPRESSO, there are fewer irreducible k-points generated by our method than Monkhorst-Pack grids generated using a more conventional method, significantly reducing the calculation time.

The hydrogen evolution reaction plays an important role for sustainable development by producing hydrogen as clean energy from water. As known, Pt is almost the best catalyst for electrochemical reaction of water splitting, but it is rare and expensive. To address this problem, we studied the hydrogen evolution reaction on CoP (101) to figure out its efficiency as catalyst, and compared its properties with Pt(100) and Pt(111) at the atomic scale using cluster expansions.

To describe the catalytic activity, the hydrogen adsorption energy depending on the coverage of adsorbed hydrogen on catalyst's surface is introduced into research. We have constructed cluster expansions of hydrogen adsorption on each of these three surfaces: Pt (100), Pt (111) and CoP (101) to identify their structures with adsorbed hydrogen at different temperature, applied potential or hydrogen chemical potential. Our work demonstrates how the activity of a surface with a variety of adsorption sites and interactions among adsorbed hydrogen molecules varies as a function of the applied potential.

Advisor: Dr. Tim Mueller

Acknowledgement

I am grateful to Prof. Tim Mueller for essential help and valuable guidance during my research in this two years. Prof. Mueller's intelligence, charming personality and enthusiasm to materials science are excellent characteristics establishing a perfect example as a respectable scholar. The time working with him is truly enjoyable and let me rethink myself trying to build a better personality. I particularly thank Pandu Wisesa and Chenyang Lee for their help. They taught me a lot and corrected my mistakes with much patience. I also acknowledge Liang Cao, Fenglin Yuan, Peter Lile, Alberto Hernandez-Valle, Chuhong Wang, Hao Gao, Yunzhe Wang and all other people I have been in contact with in Prof. Mueller's group, for fruitful conversations and comradery.

I should also thank Department of Materials Science and Engineering and Johns Hopkins University and collaboration with MIT, Brown, Emory for office of Naval Research, which providing me such an excellent environment and support during the research.

Table of Contents

List of Figures

1 Introduction

1.1 Clean Energy Resource: Hydrogen

With the development of industry and technology, the resources to produce energy have occupied a large proportion. Fossil fuels as the most common and traditional resources are facing shortage which is a serious issue now. At the same time, the pollution produced from fossil fuels pushes people give preference to green and renewable energy to release the shortage of energy and environmental problems.

At this time, we found out hydrogen turns out to be one of the most promising energy carrier, because it is clean, and has high energy density. It can be produced easily and economically via waters splitting. Therefore, a study on the water splitting is crucial and especially finding efficient catalysts for oxygen evolution reaction (OER) and the hydrogen evolution reaction (HER) has great importance for the development of solar water splitting system for the commercial production of the hydrogen fuel which generates molecular hydrogen through the electrochemical reduction of water, underpins many clean-energy technologies and therefore reduce the demand for fossil fuels.

1.2 Transition Metal Phosphides for hydrogen evolution reaction

To make the reaction more effective, we have to introduce catalysts into reaction. From the experimental research from previous research in the published papers [1,2,3], Pt is shown to be almost the best catalyst for reaction. But it is rare and expensive, we should try to find substitute which may has potential effect as catalysts. The transition metal phosphides combine the physical properties of ceramic compounds, like strength and hardness, with the transport properties of metals, such as thermal and electrical conductivity, which makes them to be an increasing potential class of catalysts. They show excellent catalytic properties in hydrogen evolution reaction [4,5,6]. From paper [7] published by Nørskov's group, we found that the transition metal phosphides (TMPs) with Fe-Co alloy have emerged as active HER catalysts with good performance and stability from both experiments and computational studies.

In this paper, we will explore catalytic behavior of transition metal phosphate - CoP and compare its efficiency with Pt.

2 Theoretical backgrounds and methods

2.1 Density functional theory

Density functional theory [8,9] is a kind of method for studying the electronic structure of multi-electron systems through electron density.

2.1.1 Hohenberg-Kohn's theorems

Hohenberg-Kohn's theorem is proposed in 1964 by P. Hohenberg and W. Kohn. On the basis of the theory of non-uniform electron gas, they brought up two basic theorems which laid the foundation of density functional theory[10].

Hnhenberg-Kohn first theorem [10]: For a single electron density $\mathrm{n(r)}$, there is only one unique corresponding ground state external potential energy term $V_{ext}(r)$ and ground state total energy E. Then the ground state energy of Schrödinger equation can be expressed using electron density:

$$E[n(r)] = \Psi^{*}\left(\hat{T} + \hat{V} + \hat{U}\right)\Psi \qquad (1)$$

The degree of freedom of the Schrödinger equation has dropped from $3N$ to 3, where N is the number of atoms in the system, which greatly reduces the computational complexity.

The first theorem of Hohenberg-Kohn can be proved by counterevidence[11]. Supposed that there are two different sets of wave functions for the same electron density, a set of ground state energies can be derived as following according to the variation principle:

$$E_1^0 < \Psi_2^* \widehat{H_1} \Psi_2 = \Psi_2^* \widehat{H_2} \Psi_2 + \Psi_2^* \left(\widehat{H_1} - \widehat{H_2}\right) \Psi_2 = E_2^0 + \int n(r) \left[V_{ext,1}(r) - V_{ext,2}(r) \right] dr \qquad (2)$$

For the same electron density, the kinetic energy terms and the interaction terms must be the same. Then we can get:

$$E_1^0 + E_2^0 < E_1^0 + E_2^0 \qquad (3)$$

which is obviously wrong. Uniqueness can be testified.

The second theorem of Hohenberg-Kohn[10]: The electron density that can obtain the lowest energy is the electron density of the ground state. It can be regarded as the variation principle of

the Hohenberg-Kohn theorem and can be derived from the variation principle of the Schrödinger equation.

2.1.2 Density Function Theory

According to the Hohenberg-Kohn theorem, the wave function Ψ, energy E and other series of quantum mechanical properties can be described by functions of the electron density[8,9]. The electron density can be written as:

$$n = n(r) = N \int dr_1 \dots \int dr_N \Psi^*(r_1, r_2 \dots r_N)\Psi(r_1, r_2 \dots r_N) \qquad (4)$$

The total energy of the system can be written as following:

$$E(n) = \Psi^*(n)\hat{H}\Psi(n) = \Psi^*(n)\left(\hat{T} + \hat{V} + \hat{U}\right)\Psi(n) \qquad (5)$$

There are only three degrees of freedom in the current equation, which makes it possible for us to analyze the larger system. The next step is to approximate and simplify the three terms of total energy: kinetic energy, external potential energy and interaction[9].

For the kinetics term, we ignore the interaction between the electrons and suppose that the electrons only exist in a single electron orbit. Then the kinetic energy term can be written as

$$T(n) \approx T_s(n) = -\frac{\hbar^2}{2m}\sum_{i=1}^{N}\int d^3r\phi_i^*(r)\nabla^2\phi_i(r) \qquad (6)$$

$\phi_i(r)$ is the i_{th} single electron orbit; $T_s(n)$ is the sum of kinetic energy of all these non-interacting single electron orbitals[9]. For the external potential terms, according to the Bonn-Oppenheimer approximation the function of external potentials can be expressed:

$$V(n) = \int V(r)n(r)dr \qquad (7)$$

For the interaction terms, we can use the Thomas-Fermi model to arrive at a very intuitive approximation:

$$U(n) \approx U_H(n) = \frac{e^2}{2}\int dr \int r' \frac{n(r)n(r)'}{|r-r'|} \qquad (8)$$

It is not difficult to see that the form of $U_H(n)$ is very close to the classical Coulomb interaction expression[9]. Adding the above three items to get an approximate energy value:

$$E_{approx}(n) = T_s(n) + U_H(n) + V(n) \qquad (9)$$

We call its error exchange-correlation energy:

$$\varepsilon = (T - T_s) + (U - U_H) = E_{XC} \qquad (10)$$

The exact total energy of the system is:

$$E(n) = T_s(n) + U_H(n) + V(n) + E_{XC}(n) \qquad (11)$$

The above is the framework of the density functional theory.

2.1.3 Exchange-related functions

The core method of density function theory is to obtain an easy-to-calculate but inaccurate result by approximation and separate all errors into another term for analysis. In this section, we will introduce several approximate methods which are often used in computational simulation of materials for calculating exchange-correlated energy.

2.1.3.1 Local Density Approximation(LDA)

LDA is the most concise exchange-correlation function and was proposed very early. In LDA method, we can approximate the E_{XC} at each point in the space is only related to the local electron density at that point[12]. The general expression of LDA can be written as

$$E_{XC}^{LDA}(n) = \int dr f(n) \qquad (12).$$

we can discuss the exchange part (X term) and the correlated part (C term) of E_{XC} separately[13]:

$$E_{XC}^{LDA} = E_X^{LDA} + E_C^{LDA} \qquad (13).$$

1) For the exchange part where the local density is approximated, a homogenous or uniform electron gas is usually used, which means that the exchange energy of an atom or molecule at some point in the space is correspond to the uniform electron gas with the same electron density at this point. The exchange energy of the uniform electron gas (UEG) can be given by the Dirac function:

$$E_X^{LDA}(n) = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{1/3} \int d^3 r n(r)^{\frac{4}{3}} \qquad (14)$$

2) For the correlated part, there is different in approximation between the high density uniform electron gas and low density one. Under high-density condition, the correlation energy at a point in space can be written as

$$\epsilon_C = A ln(r_s) + C + O(r_s) \qquad (15),$$

where r_s is Wigner-Seitz radius, $r_s = \frac{1}{a_0}(\frac{3}{4\pi n})^{-3}$; $a_0 r_s$ is the radius of a sphere containing $1e$ in UEG; $A \approx 0.0311$; $C \approx -0.048$. Under low-density conditions, the correlation energy at some point in space can be written as

$$\epsilon_C = \frac{1}{2}\left(\frac{g_0}{r_s} + \frac{g_1}{r_s^{\frac{3}{2}}} + \cdots\right) \qquad (16),$$

$g_0 \approx -0.884,\ g_1 = 3.$

The LDA was proposed very early and widely used in materials science research. In general, LDA is accurate to give good estimation of structural and elastic properties. However, it has several disadvantages, such like overestimation of binding energy, underestimation of reaction activation energy, excessive preference for high-spin structures, and miscalculation of phase stability[12].

2.1.3.2 General-Gradient Approximations(GGA)

LDA does not perform well in systems where the electron density changes rapidly. An easy way to improve it is to include one step gradient $n(r)$. So we can get the generalized gradient approximation, GGA[14,15]. The general expression of GGA can be written as[14]:

$$E_{XC}^{GGA}(n) = \int drf(n, \nabla n) \qquad (17).$$

Compared to LDA, GGA has several advantages: more accurate atomic and molecular energies, corrected over binding and more accurate reaction activation energy (but usually still too low)[14,15]. However, GGA is not always better than LDA. GGA functional usually give low bond energy and high lattice parameters.

2.1.4 The significance of density function theory

There is no doubt that the density functional theory is of great significance. Its proposal makes it possible to calculate the electronic structure of a larger system and gives more accurate results in most cases.

2.2 K-points Grid Server

In computational area, the calculation of many properties of materials needs the evaluation of integrals over the Brillouin zone in reciprocal space. These integrals are typically approximated using a discrete set of points, which we called k-points. The more number of k-points used in the same computational calculation, the more accuracy the calculation will obtain. But at the same time, the cost of time will be increased. To minimize the cost of Brillouin zone integration while maintaining sufficient accuracy, several methods have been developed for how to select k-points for Brillouin zone integration. In the Figure 2.2.1, there are three different main methods people usually use for how to generate k-points. In the figure, we can get that the most common used method is developed by Monkhorst and Pack [16]. Moreno-Soler approach tried to minimize the number of symmetrically irreducible k-points in the grid only by selecting a minimum permissible

distance between lattice points in the real space r_{min}. But it is less used than Monkhorst-Pack approach because it searches all possible super lattices in which lattice points are r_{min} to find out the fewest number of irreducible k-points for the calculation, which cost huge computational time[17,18]. To address this problem, our group developed an implement called "K-point server" to provide the fewest irreducible k-point grids when r_{min} is set[19] by using informatics.

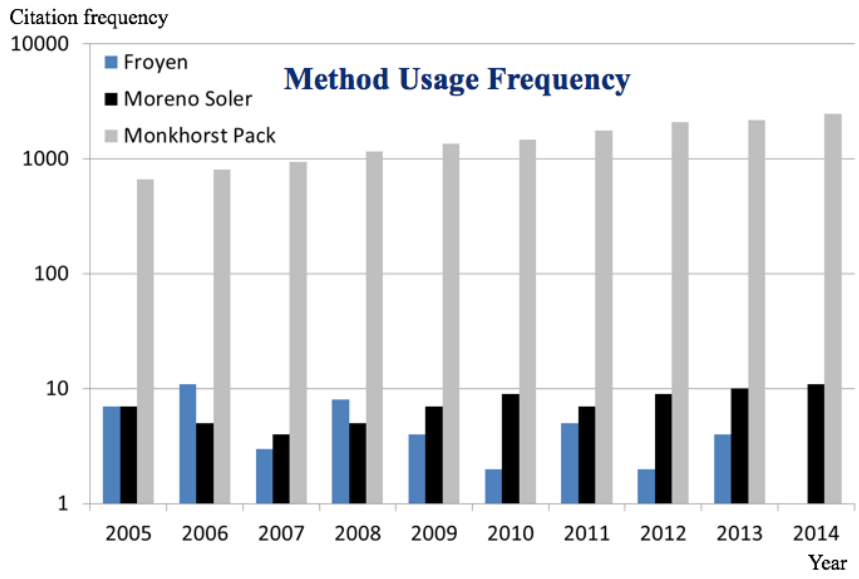


Figure 2.1 the citation frequency of different methods to generate K-point grids

Our group developed a method for rapidly generating efficient k-point grids for Brillouin zone integration by using a database of pre-calculated grids. In the method, lots of benchmarking results have been added into the database to build the k-points server, which is aimed to provide

efficient k-point grids for Brillouin zone integration. Our group use informatics to create a database of generalized k-point grids that can be rapidly searched to find the grid with the fewest irreducible k-points for which the points on the corresponding super lattice are separated by a distance of at least r_{min}. In the method, there is only one parameter, r_{min}, that needs to be set by the user. We have tried to consider all kinds of symmetries to reduce the number of irreducible k-points for calculation.

After randomly selecting different materials, our group have done various calculations using k-points generated from our server and compared with the results produced by conventional method. The results show that the grids generated by our method have much fewer irreducible k-points than Monkhorst-Pack grids generated using a more conventional method, which significantly accelerating the calculation of properties of crystalline materials, while maintaining the calculation accuracy.

From the data analysis in our group, the improvement is distinguished [19]. For Γ-centered grids, we observed average speedup of about 50–100% compared to a more conventional approach, and for shifted grids the speedup is even greater.

To encourage others to generate k-point grids using our method more conveniently, our group have constructed a free and publicly available k-point grid server that provides access to our database of generalized k- point grids. This script of server and instructions for the use of this server can be found at our group website. There is no doubt that the k-points server is a breaking through process and leads to significant acceleration of calculations on crystalline materials. However, we only generate grids in a format suitable for use with VASP at present, and we will be building interfaces to other common software packages. In the results and discussion part, we will do benchmarking calculations and discuss the application and efficiency of k-points server in Quantum ESPRESSO. At the same time, we also try to improve the quality of our database by adding additional grids, including larger grids and grids representing non-centrosymmetric space groups.

2.3 Cluster expansion

In the area of computational design of materials, the properties of crystalline materials are valued by the structure of its materials and interaction and position of atoms. In that way, a series

of functions are introduced in to connect the relationship between the fixed set of sites of crystalline materials and its various properties, which is called cluster expansion.

We define different sites as j_{th} site by s_j and the set of all such site variables by s. At each site, a single-variable site basis of functions Θ_{b_j} is defined standing for the b_{th} basis function for the j_{th} site. Then the general property, F, can be parameterized as a linear combination of these basis functions,

$$F(\mathbf{s}) = \sum_{\mathbf{b}} V_{\mathbf{b}} \prod_j \Theta_{\mathbf{b_j}}(\mathbf{s_j}) \qquad (23)$$

where $\mathbf{b_j}$, the j_{th} element of $\mathbf{b}$, is the index of the basis function to be used at site j. The sum is over all possible sets $\mathbf{b}$. The coefficients for this expansion, $V_{\mathbf{b}}$, are referred to as effective cluster interactions (ECIs). The basis function $\prod_j \Theta_{\mathbf{b_j}}(\mathbf{s_j})$ is known as a cluster function and will be represented by

$$\Phi_{\mathbf{b}}(\mathbf{s}) = \prod_j \Theta_{\mathbf{b_j}}(\mathbf{s_j}) \qquad (24)$$

If existing symmetrically equivalent cluster functions, their corresponding ECIs must be equal. Equation 1 can thus be rewritten as

$$F(\mathbf{s}) = \sum_{\alpha} V_{\alpha} \sum_{\mathbf{b} \in \alpha} \Phi_{\mathbf{b}}(\mathbf{s}) \qquad (25)$$

where the set α represents an orbit of symmetrically equivalent cluster functions and the outer sum is over all such orbits. The value of an extensive property may be normalized per unit of material. For such a property, Eq. 26 can be rewritten as

$$f(\mathbf{s}) = \sum_{\alpha} V_{\alpha} m_{\alpha} \varphi_{\alpha}(\mathbf{s}) \qquad (26)$$

where $f(\mathbf{s})$ is the average value of the property per formula unit and $\varphi_{\alpha}(\mathbf{s})$ is the average value of cluster functions in orbit α,

$$\varphi_{\alpha}(\mathbf{s}) = \frac{\sum_{\mathbf{b} \in \alpha} \Phi_{\mathbf{b}}(\mathbf{s})}{N_{\alpha}} \qquad (27)$$

The multiplicity, m_{α}, is an integer that represents the number of cluster functions in orbit α per formula unit. If all the ECIs were known, Eq. 26 could be used to exactly calculate the normalized property value for a given material state. Actually, Eq. 26 cannot be evaluated because there are infinite sites need to be considered. This problem can be addressed through truncation of the cluster expansions [20].

To construct an accurate cluster expansion with limited training data, proper values of effective cluster interaction have to be introduced into the system. We express the training data as a vector of output values, $\mathbf{y}$, and matrix of input values, X. The i_{th} element of $\mathbf{y}$ is the property value for the i_{th} element in the training set and the elements of X are given by

$$X_{i\alpha} = \varphi_{\alpha}(\boldsymbol{s_i}) \quad (28).$$

To make sure X is finite, cluster functions for which the ECIs are likely to be negligible are excluded. The probability density for the optimal ECI, given the training data, is expressed as the conditional probability distribution $P(\mathbf{v}|X,\mathbf{y})$, where the variable $\mathbf{v}$ is defined over possible ECI values.

The key method of this approach is Bayes' theorem[21] to find the ECI that maximize $P(\mathbf{v}|X,\mathbf{y})$, which can be written as

$$P(\mathbf{v}|X,\mathbf{y}) = \frac{P(\mathbf{y}|\mathbf{v},X)P(\mathbf{v}|X)}{P(\mathbf{y}|X)} \qquad (29)$$

The key to the application of Bayes' theorem is the establishment of the prior probability distribution, $P(\mathbf{v}|X)$. The prior probability distribution represents a well performed conjecture of

ECI values before we have calculated property values for the training data. In a Bayesian cluster expansion, the predicted ECI can be expressed by

$$\mathbf{v} = (\mathbf{X^T X} + \mathbf{\Lambda})^{-1}\mathbf{X^T y} \quad (30)$$

where $\mathbf{\Lambda}^{-1}\sigma^2$ is the covariance matrix for a multivariate Gaussian prior distribution for ECI values.

There are three main physical intuition considered:

1) property predictions should be close to those predicted by some simple model.

2) The greater the number of sites in the cluster, and the greater the distance between sites, the smaller the ECI should be.

3) ECI for similar cluster functions should have similar values.

Above all, we can derive a maximum like-hood estimate for $\vec{\boldsymbol{V}}$,

$$\hat{\vec{\boldsymbol{V}}} = (\mathbf{X^T W X} + \mathbf{\Lambda})^{-1}\mathbf{X^T W y} \quad (31)$$

where $\vec{\widehat{\boldsymbol{V}}}$ are the estimated ECI, $\mathbf{W}$ is a diagonal weight matrix, and $\boldsymbol{\Lambda}$ is a matrix with elements given by

$$\Lambda_{\alpha\alpha} = \frac{\sigma^2}{\sigma_\alpha^2} + \sum_{\beta|\beta\neq\alpha} \frac{\sigma^2}{\sigma_{\alpha\beta}^2} \qquad (32)$$

$$\Lambda_{\alpha\beta} = \Lambda_{\beta\alpha} = -\frac{\sigma^2}{\sigma_{\alpha\beta}^2} \qquad (33)$$

where σ^2 is an unknown constant. The regularization matrix is equivalent to using the following prior probability distribution:

$$P(\mathbf{v}|X) \propto e^{-v^{-T}\Lambda \mathbf{v}/2} \qquad (34)$$

If we limit $\lambda_{\alpha\beta} = \frac{\sigma^2}{\sigma_{\alpha\beta}^2} = 0, \lambda_\alpha = \frac{\sigma^2}{\sigma_\alpha^2}$ ranges from zero to infinite, a cross-validation (CV) method is added. In the calculations we did below, the CV error is pretty low, which means the cluster expansion we built is suitable for calculations.

2.4 Monte Carlo Simulation

Monte Carlo simulation is a statistical method used to simulate large amounts of data[22]. When the problem to be solved is the probability of occurrence of an event, or the expected value

of a random variable, they can use a "trial" method to obtain the frequency of such events or the average of the random variables. Use them as solutions to the problem. This is the basic idea of the Monte Carlo method.

In the Monte Carlo method, people use mathematical methods to grasp the number of geometry and geometric features of the motion of things, and simulate them by means of mathematical methods, that is, a digital simulation experiment. It is based on a probabilistic model and calculates the results of simulation experiments, an approximate solution to the problem, according to the process described in this model. Monte Carlo problem solving can be attributed to three main steps: constructing or describing probabilistic processes; sampling from known probability distributions; establishing various estimators.

The Monte Carlo simulation is performed on a computer simulation project for thousands of times, and each input is randomly selected input value. Since each input is often an estimation interval itself, the computer model randomly selects any value in each interval of the input, and finally obtains a cumulative probability distribution through a large number of thousands or even millions of simulations.

The Metropolis Monte Carlo method[23] is a computational algorithm for generating a set of N configurations of the system $\xi_1,\ \xi_2,\ \xi_3, \dots ,\ \xi_N$. Then we can get

$$\lim_{N\to\infty} \frac{N_\xi}{N} = P(\xi) \quad (35)$$

where $P(\xi)$ is a given probability distribution and N_ξ is the number of configurations ξ

The Metropolis Monte Carlo algorithm can be described as follows:

1) Pick a configuration ξ_n (the initial configuration can be *any* configuration of the system.

2) Pick a *trial configuration* ξ_i and compute the probability ratio

$$R = \frac{P(\xi_i)}{P(\xi_n)} \quad (36)$$

Pick a random number P with value between 0 and 1. Make $\xi_{n+1} = \xi_i$ if $p < R$. Otherwise, make $\xi_{n+1} = \xi_n$.

3) Go to step (2) replacing ξ_n by ξ_{n+1}.

Step (3) has to be repeated by N times, where N is a sufficiently large number.

In this paper, we use Metropolis monte carlo method to assume the hydrogen coverage and the possibility for hydrogen atom locating on different active sites.

2.5 Packages used in the calculation

2.5.1 VASP

VASP, which full name is Vienna Ab-initio Simulation Package, is a software package developed by the Hafner Group of the University of Vienna[24] for the calculation of electronic structures, quantum mechanics and molecular dynamics simulation. We can use VASP to obtain the electronic states and energies of the system by approximating the Schrödinger equation. It can solve the Kohn-Sham equation within the framework of the density functional theory (DFT) and is also possible to solve the Roothaan equation under Hartree-Fock (HF) approximation.

2.5.2 Quantum ESPRESSO

Quantum ESPRESSO software package[25] is a first-principles calculation software[26] based on the density functional theory applying of the plane-wave basis set and the pseudopotential method. It includes two major modules: PWscf and CPMD. In addition, there are two auxiliary graphical interface modules for input parameter setting and generation of potentials.

It can compute Fermi surfaces (metals), electro-acoustic coupling and superconducting properties (including isotropic, anisotropic superconducting features). The advantages of Quantum ESPRESSO are functional modularity (easy to add new modules) and free open source.

3 Results and Discussion

3.1 The accuracy of K-points server in Quantum ESPRESSO

In computational material area, various properties of materials are determined by integrals over the Brillouin zone in reciprocal space. K-points are a set of discrete points to approximate these integrals. Our group developed and implemented called a k-point server to provide the fewest irreducible k-point grids for a required level of accuracy. Before this paper, our group had had implemented our grids into VASP and found out the computational cost of using this method is twice as low as the conventional methods for generating Monkhorst-Pack grids. In this part, we try to test the efficiency of K-points Server in Quantum ESPRESSO to see if it can reduce the computing space and save time while maintaining accuracy.

To test for the effectiveness of our K-points server, we used density functional theory (DFT) [27,28] as implemented in Quantum ESPRESSO to calculate the converged energies of randomly selected materials from the Inorganic Crystal Structure Database (ICSD) [29]. The Perdew−Burke−Ernzerhof (PBE) exchange-correlation functional was used for all DFT calculations.

First, we pre-relaxed all materials we will test, and done a static run on the relaxed structures for benchmarking. To determine the number of k-points required to calculate a converged energy value, we generated k-point grids for 20 different values of r_{min} for each material from $r_{min} = 100\ \dot{A}$ and incrementally reducing r_{min} by a factor of $2^{1/6}$ until we reached $r_{min} = 12.5\ \dot{A}$ to corresponding to the k-points automatically generated by Quantum ESPRESSO from 20x20x20 to 1x1x1 for comparsion. After calculation, we compared the number of irreducible k-points and CPU time used when calculations reach same converged energy value for both Γ-centered and shifted grids. Γ-centered grid contains the grid at gamma point which has high symmetry than shifted grid which does not include gamma point. Γ-centered grid is supposed to use more k-points for calculation, because the Gamma point is not symmetrically equivalent to any other point.

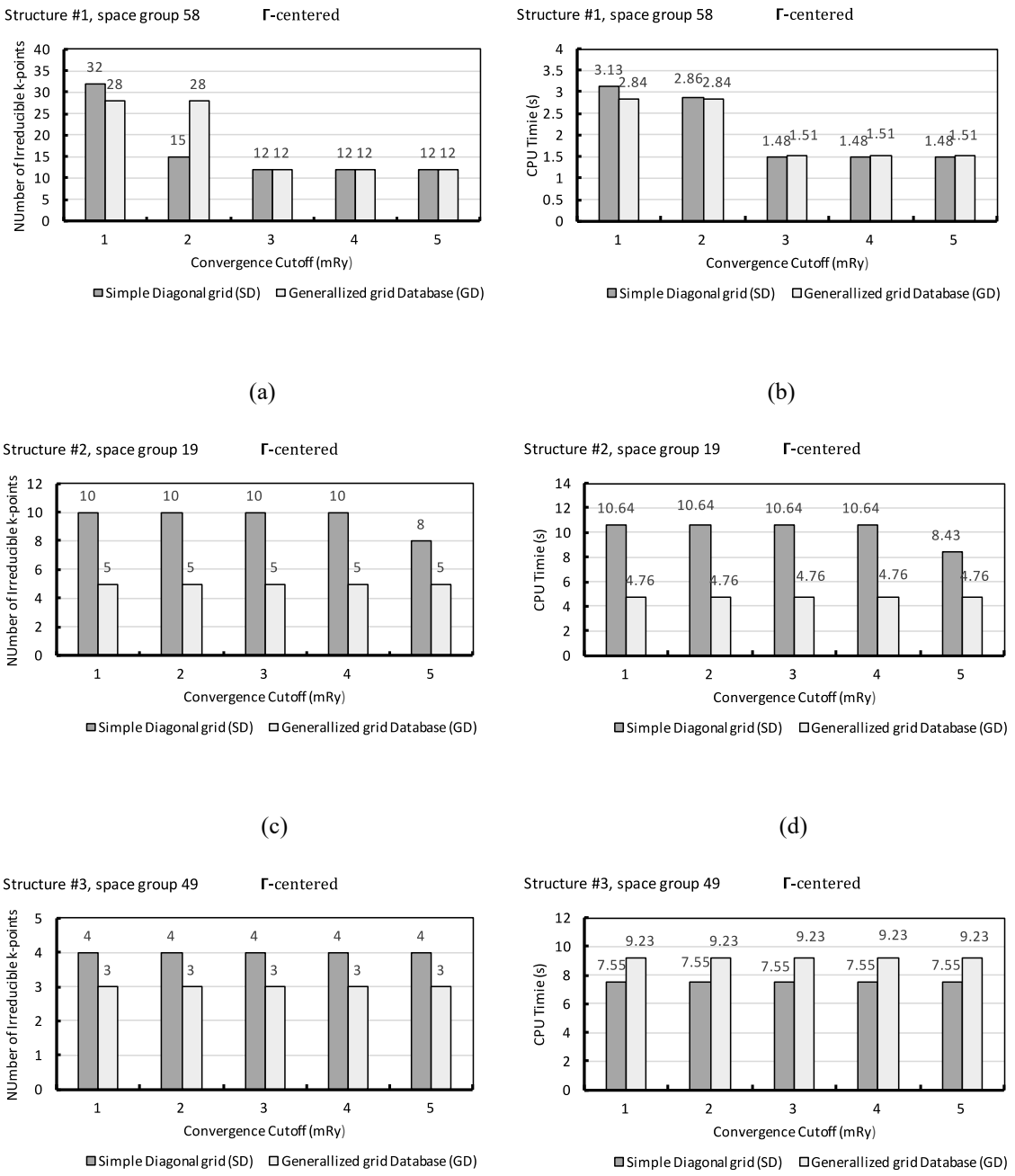
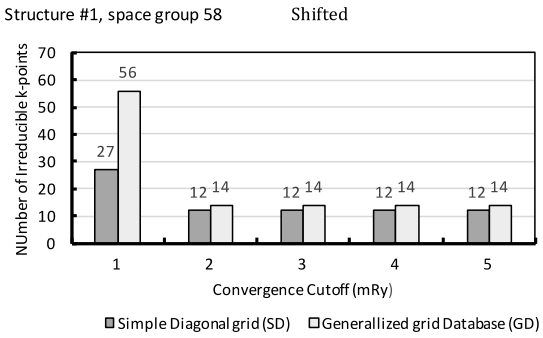


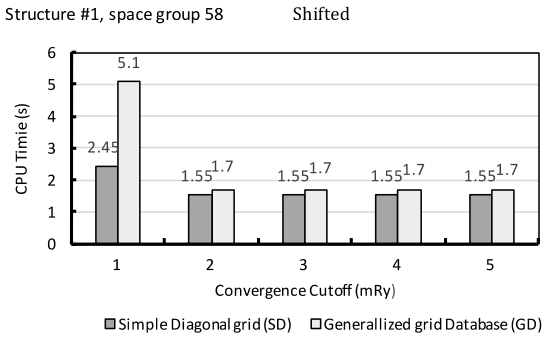
Figure 3.1 The number of irreducible k-points (a) (c) (e) and CPU time (b) (d) (f) used when calculations reach same converged energy value for Γ-centered grids on structures in space group: 58,19,49.

In the Figure 3.1, three different materials from different space group are shown to present the

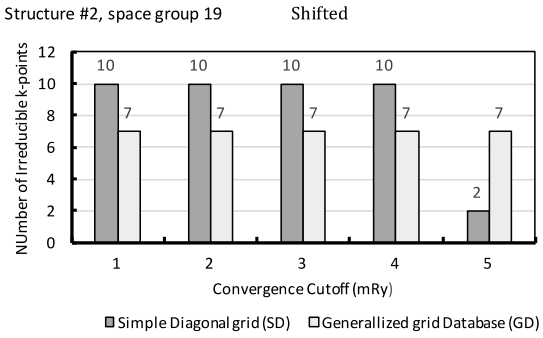
efficiency of K-points server we use. It shows both the number of k-points and CPU time used when calculations reach same converged energy value for Γ-centered grids. From the Figure 3.1, for convergence of the energy within 1 $mRy/atom$, the Simple Diagonal (SD) method required on average 1.49 times as many irreducible k-points as the Generalized grid Database (GD) method for Γ-centered grids and takes around 1.27 times CPU time to finish the calculation. Here Simple Diagonal method means automatic k-point generation scheme used in the Quantum ESPRESSO [25,26] and the Generalized grid Database (GD) method is the method used in our K-point Server.



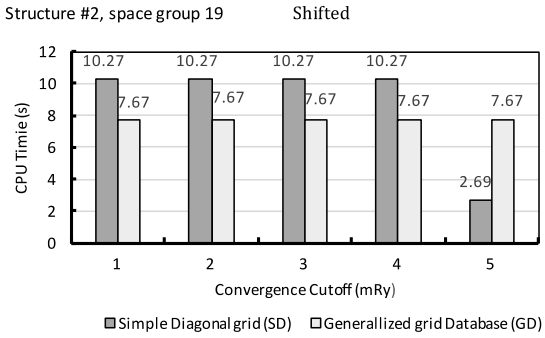
(a)



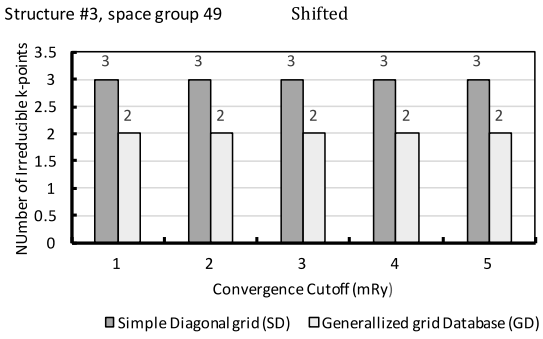
(b)



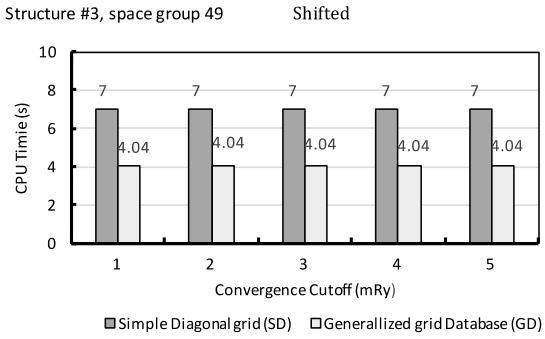
(c)



(d)



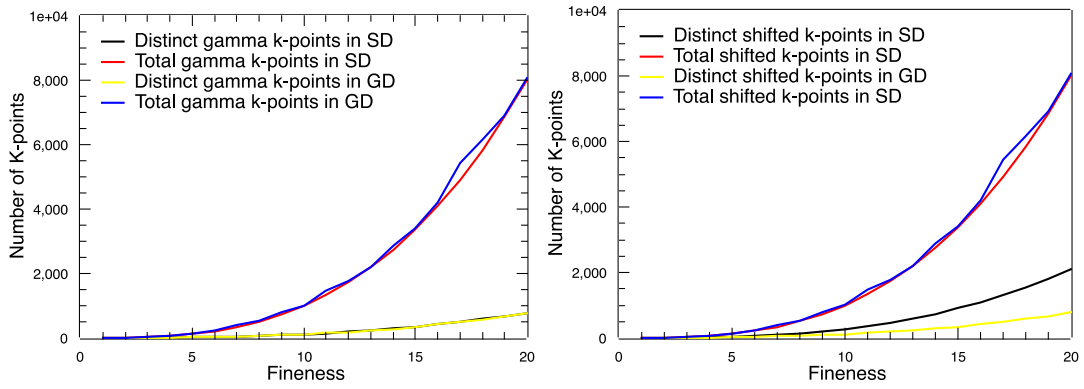
(e)



(f)

Figure 3.2 The number of irreducible k-points (a) (c) (e) and CPU time (b) (d) (f) used when calculations reach same converged energy value for shifted grids on structures in space group: 58,19,49.

Similar with the Figure 3.1, in the Figure 3.2, the number of k-points and CPU time used when calculations reach same converged energy value for shifted grids are compared. From the Figure 3.2, for convergence of the energy within 1 $mRy/atom$, the SD method required on average 1.14 times as many irreducible k-points as the GD method for shifted grids and cost 1.17 times CPU time compared with using GD method.



(a) (b)

Figure 3.3 The relationship between the number of irreducible and total k-points used with fineness for Γ-centered (a) and shifted (b) grids on structure in space group: 49

Then we chose the material structure in space group 49 to do specific research on the number of k-points used of Brillion zone in different fineness. In the Figure 3.3, difference can be easily noticed in the plot. There are more irreducible k-points used in SD method compared with GD method applied in k-point server, and this difference is more distinguishing when using shifted k-

point grids for calculation. On average, the SD method required on average 1.03 times as many distinct k-points as the GD method for Γ-centered grids and 2.32 times for shifted grids.

The relative advantage of GD methods over the SD method increases as the convergence criterion is tightened. This is primarily because the benefit of using highly symmetric grids, in terms of the reduction in the total number of irreducible k-points, is greater for grids that have a large number of total k-points. At the same time, the trends of computational time for the entire batch of materials are similar to those for the average number of irreducible k-points. The main reason is because computational time usually scales as the number of irreducible k-points. Once gamma k-point, which almost has the highest symmetry, is removed, the number of irreducible k-points is largely increased in SD method to make up the losing symmetry, while the GD method determines the k-points grids differently which causes that the number of irreducible k-points in GD method does not have a dramatic change.

3.2 Hydrogen Evolution Reaction

HER is a stepwise chemical reaction that involves three elementary steps: Volmer, Heyrovsky, and Tafel reaction [30]. Those steps can be expressed as

$$H^+ + e^- + * \rightarrow H^* \text{ (Volmer)} \qquad (37)$$

$$H^+ + e^- + H^* \rightarrow H_2 + * \text{ (Heyrovsky)} \qquad (38)$$

$$H^* + H^* \rightarrow H_2 + 2^* \text{ (Tafel)} \qquad (39)$$

In the above equation, $*$ indicates the active site, H^* indicates one hydrogen adsorbed on the active site. From eq. 30,31,32, we can know under acidic condition, one hydrogen molecule is produced from H^* adsorbed on the surface of the catalyst. And H^* is formed from an electron with an adsorbed proton. Therefore, the hydrogen binding energy is considered as the descriptor for the HER activity [31,32].

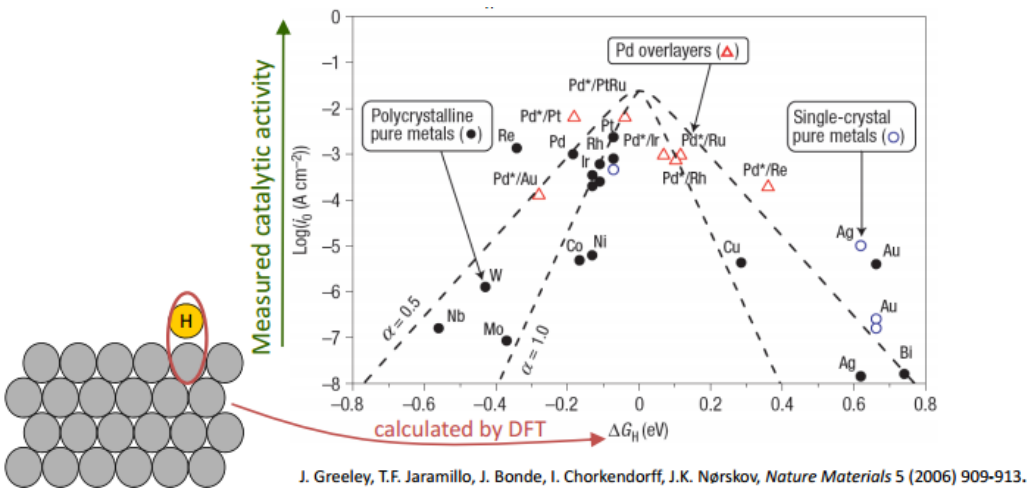
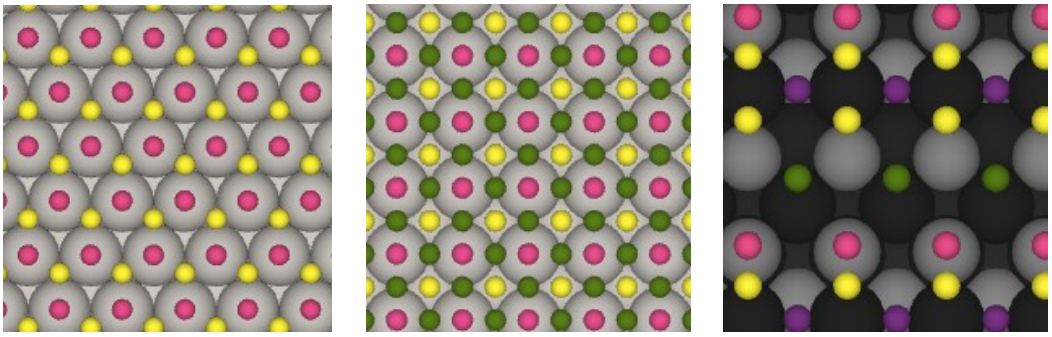


Figure 3.4 Volcano plot for the HER for various pure metals and metal overlayers [33].

As shown in the Figure 3.4 cited from Nørskov's group [33], neither too strong nor too weak binding would be preferred for the whole reaction because strong or weak binding leads to either difficulty in removing the final product or poor adsorption of the reactant. It shows that the hydrogen evolution activity tends to be a monotonic decreasing when the absolute value of hydrogen binding increases.



(a) (b) (c)

Figure 3.5 different types of adsorption sites on (a) Pt (111), (b) Pt (100) and (c) CoP (101)

First, we cut the slab and set primitive structures of Pt (100), Pt (111) and CoP (101) surfaces. For primitive structures, there are three different active sites: bridge, hollow and top for adsorbed hydrogen on Pt (100) surface. For Pt (111) surface, there is FCC and top sites, and for CoP, there is Co-top, Co-Co, Co-P and P-top sites, which is shown in Figure 3.5. In the Figure 3.5 (a)(b) grey atoms are Pt, and in the Figure 3.5 (c) black atoms are Co and grey atoms are P. In Figure 3.5 (a),

pink atoms are Pt-top adsorption sites, and yellow atoms are Pt-FCC adsorption sites. In Figure 3.5 (b), pink atoms are Pt-top adsorption sites, green atoms are Pt-bridge adsorption sites and Pt-hollow adsorption sites are yellow. In Figure 3.5 (c), pink atoms are P-top adsorption sites, green atoms are Co-top adsorption sites, purple atoms stand for Co-Co-bridge adsorption sites and Co-P-bridge adsorption sites are yellow.

After that, we constructed different sizes of structures with different coverage number of hydrogen as training structures for the cluster expansion building. In this step, the more training structures we used, the better assumption of cluster expansion we will build. To build cluster expansion, we performed DFT calculations of training structures by using VASP package.

The hydrogen adsorption energy has been shown to be a valuable descriptor for HER activity. Here we represent the adsorption energy as the potential energy of hydrogen adsorption, ΔE_H, calculated as

$$\Delta E_H = E(slab + H^*) - E(slab) - \frac{1}{2}E(H_2) \quad (40)$$

$E(slab + H^*)$ is the total DFT energy of the slab with an adsorbed hydrogen atom, $E(slab)$ is the DFT energy of the clean slab, and $E(H_2)$ is the DFT energy of gas-phase H_2. By this definition, lower ΔE_H indicates stronger binding between hydrogen and the Pt/ CoP surface, while larger ΔE_H indicates weaker hydrogen binding. The free energy with adsorption of H is estimated at applied potential $U = 0V$ according to

$$\Delta G(0) = \Delta E_H + \Delta E_{ZPE} + T\Delta S \quad (41)$$

where ΔE_H is the hydrogen adsorption energy calculated using DFT. The zero-point energy corrections ΔE_{ZPE} is the zero point energy of adsorbed hydrogen atom on the active site using MD. Because $\Delta S = -\frac{1}{2}S_0(H_2)$ is the gas phase entropy, which is always same for different site, we can assume that

$$\Delta G(0) = \Delta E_H + \Delta E_{ZPE} + T\Delta S \quad (42)$$

The effect of an applied potential U can be included into the reaction

$$\Delta G(U) = \Delta G(0) - neU \quad (43)$$

where n is the number of electrons involved in the reaction.

From the above explanation, we put training structures with their relative hydrogen adsorption energy into cluster expansion to get ECIs trying to express the energy of the system using a cluster expansion. All density functional theory calculations were performed with the VASP. The Perdew−Burke− Ernzerhof (PBE) exchange-correlation functional was used for all DFT calculations.

For hydrogen atom, spin will have two different options, -1 and $+1$, where $s_i \in \{-1, +1\}$. We suppose that there are n adsorption sites on the surface and set spin for them: the spin of hydrogen atom is $+1$, and the spin of vacancy is -1. According to Section 2.3, the hydrogen adsorption energy of this surface can be expressed as following

$$E(s) = V_0 + \sum_{cluster} V_{cluster} \prod_{i \in cluster} s_i \quad (44)$$

V_0 and $V_{cluster}$ are constant. According to eq (44) and hydrogen adsorption energies of training structures, we can build cluster expansion to predict how the partial coverage and local atomic disorder can affect the adsorption energy of hydrogen. We have generated cluster expansions for

hydrogen on Pt (100), Pt (111) and CoP (101) surfaces depending on temperature and applied potential which are shown below. The scale rang of hydrogen adsorption energy is from $-0.381eV$ to $0.381eV$.

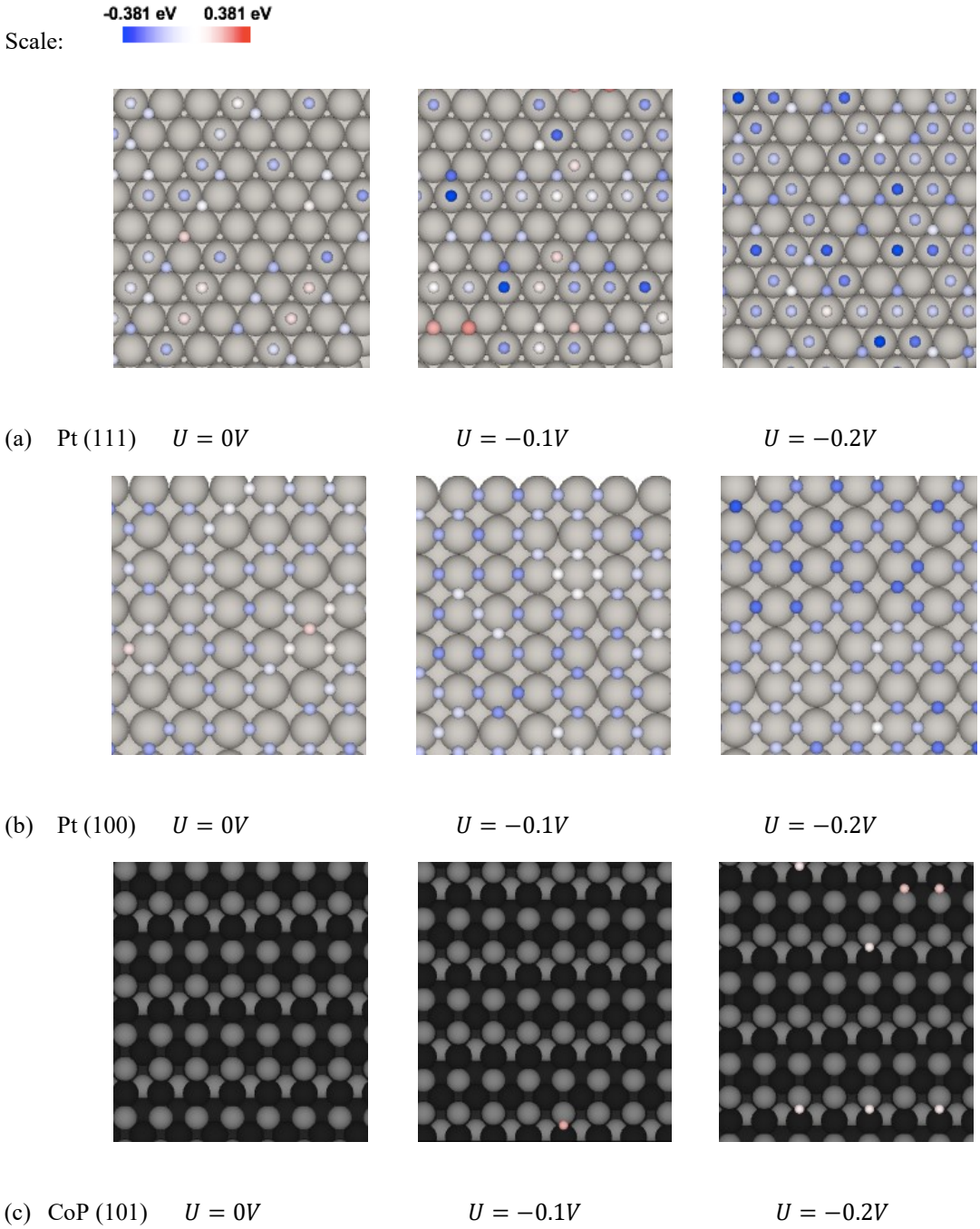


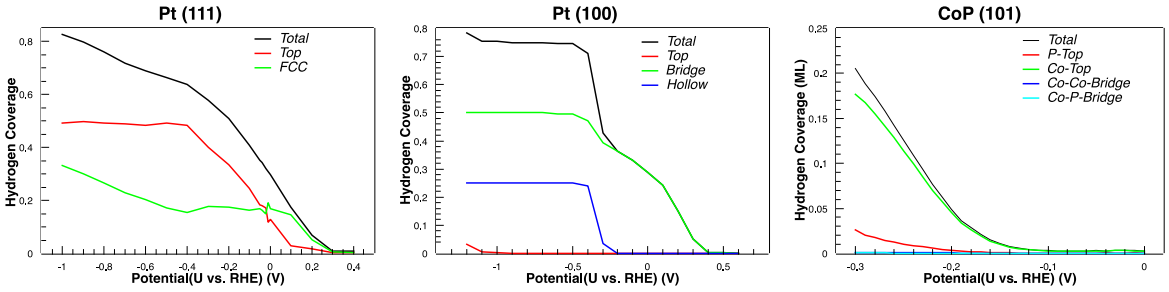
Figure 3.6 hydrogen adsorbed on (a) Pt(111) (b) Pt(100) and (c) CoP(101) when applied potential is 0V, 0.1 V and 0.2V

In the Figure 3.6 (a)(b) grey atoms are Pt, and in the Figure 3.6 (c) black atoms are Co and grey atoms are P. Other atoms colored from blue to red are the hydrogen atoms on the active sites. The

redder color is, the higher adsorption energy the hydrogen atom has. We constructed the surface of Pt (100), Pt (111) and CoP (101) under different applied potential. In the computational hydrogen electrode model, the chemical potential of a proton-electron pair is equal to half of the gas hydrogen chemical potential at $0\,V$ vs. the reversible hydrogen electrode(RHE)

$$\mu\,(H^+) + \mu(e^-) = \frac{1}{2}\,\mu(H_2) - eU. \qquad (45)$$

From the Figure 3.6, we can know that Pt is as known the best catalyst for HER and requires very small applied potential. For CoP (101), when applying a small applied potential on the (101) surface, there will exist some active sites for hydrogen which makes it possible to be a potential catalyst for HER.



(a) (b) (c)

Figure 3.7 the relationship between applied potential and hydrogen coverage on (a) Pt (111), (b) Pt (100) and (c) CoP (101)

Then we studied on the hydrogen coverage under the range of applied potential from $-0.3\,V$ to $0V$. Here all potentials were converted and referred to the RHE. In Figure 3.7 (a), we can see for Pt (111) surface, when applied potential is zero, there are around $0.3ML$ hydrogen adsorbed on Pt surface, and the top and FCC sites on the surface show the same importance as active sites for hydrogen. For Pt (100) surface in Figure 3.7 (b), when applied potential is zero, there are around $0.3ML$ hydrogen adsorbed on Pt surface, and the bridge sites on the surface are the most active sites for hydrogen. When the applied potential goes smaller, FCC sites take more responsibility. Figure 3.7 (a) and (b) demonstrates Platinum's good performance. And for CoP (101) in Figure 3.7(c), when applied potential is zero, we cannot see its catalyst behavior, while when the applied potential goes smaller, the hydrogen adsorbed phenomenon goes clear. What's more, the Co-top sites on the surface are the most active sites for hydrogen atom.

4 Conclusion

4.1 The accuracy of K-points server in Quantum ESPRESSO

Using informatics, the k-point server based on benchmarking data base successfully provides fewer irreducible k-points for calculation and reduces its complexity and cost in Quantum ESPRESSO. The generated k-point grids are always consistent with the symmetry of the material, which both preserves the symmetry of the system and maximizes the degree to which symmetry can be used to reduce the cost of the calculation. The grids are independent of the lattice vectors chosen to represent the real-space primitive cell. Grids generated using our method result in a significant reduction in the number of irreducible k-points required to reach a given level of convergence, resulting in large savings in computational time.

4.2 Hydrogen Evolution Reaction

From the previous research, cluster expansion has been well developed to determine how the activity of a surface with a variety of adsorption sites and interactions among adsorbed hydrogen atoms varies as a function of the applied potential. We can see when applied potential is zero, there are $0.3ML$ hydrogen adsorbed on Pt surface, which showing its good performance. And for CoP

(101), when applying a little applied potential, there will be active sites on surface for hydrogen. When applied potential is around $-0.25V$, the hydrogen adsorbed phenomenon goes clear. It is demonstrated that CoP is a potential catalyst for hydrogen evolution reaction.

Reference

[1] Nørskov, J. K. *et al*. Trends in the exchange current for hydrogen evolution. *J. Electrochem. Soc.* **152,** J23–J26 (2005).

[2] Li, Yanguang, et al. "MoS2 nanoparticles grown on graphene: an advanced catalyst for the hydrogen evolution reaction." *Journal of the American Chemical Society* 133.19 (2011): 7296-7299.

[3] Gasteiger, Hubert A., et al. "Activity benchmarks and requirements for Pt, Pt-alloy, and non-Pt oxygen reduction catalysts for PEMFCs." *Applied Catalysis B: Environmental*56.1-2 (2005): 9-35.

[4] Oyama, S. Ted. "Novel catalysts for advanced hydroprocessing: transition metal phosphides." *Journal of Catalysis* 216.1-2 (2003): 343-352.

[5] Liu, Qian, et al. "Carbon Nanotubes Decorated with CoP Nanocrystals: A Highly Active Non-Noble-Metal Nanohybrid Electrocatalyst for Hydrogen Evolution." *Angewandte Chemie*126.26 (2014): 6828-6832.

[6] Popczun, Eric J., et al. "Highly active electrocatalysis of the hydrogen evolution reaction by cobalt phosphide nanoparticles." *Angewandte Chemie* 126.21 (2014): 5531-5534.

[7] Jakob Kibsgaard, Charlie Tsai, Karen Chan, Jesse D. Benck, Jens K. Nørskov, Energy Environment Sci, 2015, 8, 3022-3029.

[8] Gross, Eberhard KU, and Reiner M. Dreizler, eds. *Density functional theory*. Vol. 337. Springer Science & Business Media, 2013.

[9] Yang, Weitao, and Paul W. Ayers. "Density-functional theory." *Computational Medicinal Chemistry for Drug Discovery*. CRC Press, 2003. 103-132.

[10] Kryachko, Eugene S. "Hohenberg-Kohn theorem." *International Journal of Quantum Chemistry* 18.4 (1980): 1029-1035.

[11] Görling, Andreas. "Density-functional theory beyond the Hohenberg-Kohn theorem." *Physical Review A* 59.5 (1999): 3359.

[12] Stampfl, C., and C. G. Van de Walle. "Density-functional calculations for III-V nitrides using the local-density approximation and the generalized gradient approximation." *Physical Review B* 59.8 (1999): 5521.

[13] Jackson, Koblar, and Mark R. Pederson. "Accurate forces in a local-orbital approach to the local-density approximation." *Physical Review B* 42.6 (1990): 3276.

[14] Perdew, John P., Kieron Burke, and Matthias Ernzerhof. "Generalized gradient approximation made simple." *Physical review letters* 77.18 (1996): 3865.

[15] Perdew, John P., et al. "Atoms, molecules, solids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation." *Physical Review B* 46.11 (1992): 6671.

[16] H. J. Monkhorst and J. D. Pack, Phys. Rev. B 13, 5188 (1976).

[17] Soler, José M., et al. "The SIESTA method for ab initio order-N materials simulation." *Journal of Physics: Condensed Matter*14.11 (2002): 2745.

[18] Gonze, Xavier, et al. "ABINIT: First-principles approach to material and nanosystem properties." *Computer Physics Communications* 180.12 (2009): 2582-2615.

[19] P. Wisesa, K. A. McGill, and T. Mueller, Physical Review B 93, 155109 (2016).

[20] Mueller, Tim, and Gerbrand Ceder. "Bayesian approach to cluster expansions." *Physical Review B* 80.2 (2009): 024103.

[21] Mueller, Tim, and Gerbrand Ceder. "Exact expressions for structure selection in cluster expansions." *Physical Review B*82.18 (2010): 184107.

[22] Robert, Christian P. *Monte carlo methods*. John Wiley & Sons, Ltd, 2004.

[23] Metropolis, Nicholas, and Stanislaw Ulam. "The monte carlo method." *Journal of the American statistical association*44.247 (1949): 335-341.

[24] Kresse, G. "Software vasp, vienna, 1999; g. kresse, j. furthmüller." *Phys. Rev. B* 54.11 (1996): 169.

[25] Giannozzi, Paolo, et al. "QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials." *Journal of physics: Condensed matter* 21.39 (2009): 395502.

[26] Scandolo, Sandro, et al. "First-principles codes for computational crystallography in the Quantum-ESPRESSO package." *Zeitschrift für Kristallographie-Crystalline Materials*220.5/6 (2005): 574-579.

[27] P. Hohenberg and W. Kohn, Phys. Rev. 136, B864 (1964).

[28] W. Kohn and L. J. Sham, Phys. Rev. 140, A1133 (1965).

[29] Inorganic Crystal Structure Database, (Fiz Karlsruhe), http://www.fiz-karlsruhe.de/icsd.html.

[30] Bhardwaj, Mukesh, and R. Balasubramaniam. "Uncoupled non-linear equations method for determining kinetic parameters in case of hydrogen evolution reaction following Volmer–Heyrovsky–Tafel mechanism and Volmer–Heyrovsky mechanism." *International Journal of Hydrogen Energy* 33.9 (2008): 2178-2188.

[31] Trasatti, Sergio. "Work function, electronegativity, and electrochemical behaviour of metals: III. Electrolytic hydrogen evolution in acid solutions." *Journal of Electroanalytical Chemistry and Interfacial Electrochemistry* 39.1 (1972): 163-184.

[32] Vojvodic, Aleksandra, and Jens K. Nørskov. "New design paradigm for heterogeneous catalysts." *National Science Review* 2.2 (2015): 140-143.

[33] Greeley, Jeff, et al. "Computational high-throughput screening of electrocatalytic materials for hydrogen evolution." *Nature materials* 5.11 (2006): 909.

Curriculum Vitae

Wan Wan

Email: wwan4@jhu.edu

Address: Department of Materials Science and Engineering Johns Hopkins University, Whiting School of Engineering, Maryland Hall 138A, 3400 North Charles Street, Baltimore, MD 21218-2608

Tel: 443-248-1043

Education

Johns Hopkins University **Baltimore, United States**
Master in Materials Science and Engineering Sept.2016-May 2018

- GPA: 3.73/4.0

University of Science and Technology Beijing **Beijing, China**
Bachelor of Science in Nanomaterials and Nanotechnology Sept.2012-Jun.2016

- GPA: 3.41/4.0 Major GPA: 3.57/4.0
- Third prize of "Ren Min" scholarship Jan.2013
- First prize the USTB 16th Extracurricular Scientific Work competition Dec.2014

Research Experience

Johns Hopkins University (JHU) **Baltimore, United States**

Research Assistant Advisor: Dr. Tim Mueller Sept.2016-May 2018

- Generated K-point Grids for Quantum ESPRESSO to test K-points server built in group and cooperate to complete the Java scripts; used less K-points for density function theory (DFT) calculation to save time and cost while remaining calculation accuracy
- Collaborated with MIT, Brown, Emory for office of Naval Research; used VASP to analyze calculation of cluster expansion research for Hydrogen Evolution Reaction; developed 4 potential phosphide catalysts for hydrogen evolution reactions by calculating hydrogen-surface bonding energy using Java and Linux language; generated the possible catalysts structures via quantum Monte Carlo calculation; identified the hydrogen adsorption mechanism and optimal HER catalysts

- Taught students how to use Matlab and COMSOL to solve problems in kinetics, mechanical properties, and structure of materials as a teaching assistant
- Attended Materials Research Society 2017 Fall meeting in Boston and done a poster to present our project work.

University of Science and Technology Beijing (USTB) **Beijing, China**

Research Assistant Advisor: Dr. Qingliang Liao **Mar.2016-Jun.2016**

Preparation of MSNs Carrier for Programmed Releasing Medication

- Combined mesoporous silica nanoparticles with thermal sensitive materials to build drug carrier.
- Analyzed prepared mesoporous silica materials at different stages with FT-IR to demonstrate each step is developed as expected route during the study successfully.
- Tested the programmed releasing ability of drug carrier under conditions of thermal response through establishing models in different thermal conditions.

University of Science and Technology Beijing (USTB) **Beijing, China**

Research Assistant Advisor: Dr. Yan li **Sept.2014-Jan.2015**

Improving Photocatalytic Performance of ZnO via Synergistic Effects of Ag Nanoparticles and Grapheme Quantum Dots

- Developed a simple and "green" method for preparing the ternary photocatalyst Ag-grapheme quantum dots (GQDs)-ZnO.
- Fabricated the ternary photocatalyst consisting of wurtzite structured ZnO, face-centered cubic Ag, and graphite-structured GQDs through UV irradiation, spin coating, and subsequent annealing with an aqueous solution of GQDs both as a solvent and a reducing agent.
- Characterized the samples by X-ray diffraction (XRD), scanning electron microscopy (SEM), transmission electron microscopy (TEM), and photocatalytic tests involving the degradation of Rhodamine B.
- Discovered that the synthesized ternary photocatalyst displayed excellent visible-light photocatalytic activity and which was much higher than that of pure ZnO and binary photocatalysts.
- Co-author of the namesake paper. Published on Physical Chemistry Chemical Physics.

- Co-author of the paper titled *Redox-controllable Photoluminescent Performance of Graphene Quantum Dots by Post Redox Treatments* submitted to Nanoscale

Research Assistant Advisor: Dr. Ning Ma **Sept.2014-Dec.2014**

Nonenzymatic Glucose Sensor based on in-situ Reduction of Ni/NiO-graphene Nanocomposite

- Characterized the morphology and structure of the obtained rGO sheets by Field Emission Scanning Electron Microscope (FESEM), Raman spectrometer and X-rayphotoelectron spectroscopy (XPS) and studied the reaction parameters of the Ni/NiO-Nafion-rGO/SPE via cyclic voltammetry.
- Obtained the biosensor which has high electrochemical activity for electrocatalytic oxidation of glucose in alkaline medium and can be functional in a wide linear range of glucose from 29.9 μM to 6.44 mM (R=0.9937), with a low detection limit of 2.3 μM (S/N=3), a fast response of 2s, and a high sensitivity of 1557 μA/mM·cm-2.

Research Assistant Advisor: Dr. Junjie Qi **Oct.2013-Oct.2014**

Synthesis of Graphene via Chemical Vapor Deposition (CAD) and Construction of Optical Probe Device

- Studied the synthesis of high quality graphene film on Cu foil by atmospheric pressure chemical vapor deposition (APCVD) systematically.
- Investigated the influence of technological parameters, including the growth temperature and time, the methane flow and the hydrogen flow, on the structure of graphene films in detail and drew conclusions on the optimal growth parameters.
- Constructed an optical detector with zinc oxide nanometer array and obtained graphene films under the optimal condition.

Internship Experience

Ningbo Institute of Industrial Technology, CAS **Ningbo, China**

Research Assistant Advisor: Prof. Aiguo Wu **Jun.2015-Sept.2015**

Study on Upconversion Nanophosphors Excited by 808nm and their Application in Tumor Therapy Guided by Magnetic Resonance Imaging (MRI)

- Reviewed the latest literature on upconversion nanoparticles (UCNPs) and its applications in theranostics. Additionally, provided a recapitulation of the papers in which various

preparation methods for UCNPs were mentioned and investigations on UCNPs attached with gold nanoparticles were also described.

- Repeated the critical experiments of upconversion nanophosphors synthesis by high temperature thermal decomposition method, mesoporous silica synthesis and modification, and grafting characteristics and graft ratio investigation via high performance liquid chromatography (HPLC).
- Finished the preparation of gold nanostars and helped in cell culture and other cell experiments.
- Further investigations on the applications of gold nanostars coating on mesoporous silica is ongoing.

Beijing Zhongkexin Electronics Equipment Company **Beijing, China**
Assistant Coordinator **Jul. 2015**

- Worked in Solar Panel workshops to check product quality through electroluminescence.
- Learned the basic knowledge of refractory productions, including material choice as well as material molding, sintering, polishing and heaping bricks.

Wen Ti Bird Company **Beijing, China**
Teaching Assistant **Jul 2014-Dec 2014**

- Helped to answer questions and solve matters for high school students.

Skills

- UV irradiation, spin coating, COMSOL, VASP, XRD, SEM, TEM, FESEM, Raman spectrometer, XPS, cyclic voltammetry, high temperature thermal decomposition method, HPLC, cell culture.
- Proficient in using Microsoft Office Suite (Word, Excel and PPT), Orign8, Mathematica and Matlab, C++, Java